

Conditional entropy and Landauer principle

D. CHIUCHIÚ^{1 (a)}, M. C. DIAMANTINI^{1,2} and L. GAMMAITONI¹

¹ *NiPS Laboratory, Dipartimento di Fisica, Università degli studi di Perugia - Via Alessandro Pascoli, I-06123 Perugia, Italy*

² *INFN, sezione Perugia - I-06100 Perugia, Italy*

PACS 05.70.Ln – Nonequilibrium and irreversible thermodynamics

PACS 05.10.Gg – Stochastic analysis methods

Abstract – Landauer principle describes the minimum heat produced by an information-processing device. Recently a new term has been included in the minimum heat production: it's called conditional entropy and takes into account the microstates content of a given logic state. Usually this term is assumed zero in bistable symmetric systems thanks to the strong hypothesis used to derive Landauer principle. In this paper we show that conditional entropy can be nonzero even for bistable symmetric systems and that it can be expressed as the sum of three different terms related to the probabilistic features of the device. The contribution of the three terms to conditional entropy (and thus to minimum heat production) is then discussed for the case of bit-reset.

I. Introduction. – Landauer principle relates information theory and thermodynamics [1]: it express a quantitative relation for the minimum heat produced by a computing device during information processing. Its simplest formulation states that at least $K_b T \ln 2$ of heat has to be produced to reset one bit in binary symmetric systems. Recently, a more sophisticated formulation has been proposed [2,3] that takes into account the role of conditional entropy [4] to relate Shannon and Gibbs Entropy. In these works the role of conditional entropy has been emphasized in the presence of asymmetric bistable systems. In this paper we show that asymmetry is not a necessary condition and that conditional entropy can significantly affect the minimum heat production even for symmetric devices. To see this point, we consider a computing device as generic physical system with an Avogadro number N of degrees of freedom (DOF). Those are classified in two categories: one DOF is chosen to characterize the relevant system dynamic and the computation process, while the remaining $N - 1$ DOFs behave as a thermal bath at constant temperature T . The dynamic-relevant DOF is labelled x and, by construction, it can assume a large number of values, possibly continuum, each corresponding to one different microstate of the system. The set of possible x values, Ω , is the ensemble of all the microstates of the device. Because of the thermal bath, x fluctuates and we can define

$P(x)$, the probability density function (PDF) of x . The thermodynamic Gibbs entropy [5,6] of the system is then defined as:

$$S_G = -K_b \int_{\Omega} P(x) \ln P(x) dx \quad (1)$$

where K_b is Boltzmann constant.

To encode n different logic states in the device, Ω is divided in n non-overlapping subsets $\Omega_0 \dots \Omega_{n-1}$, each one containing the microstates consistent with the given logic state¹. The probability of assuming the i -th logic state is

$$P_i = \int_{\Omega_i} P(x) dx \quad (2)$$

and the information content of the device is described by Shannon information entropy [4,7]

$$S_S = -K_b \sum_{i=0}^{n-1} P_i \ln P_i. \quad (3)$$

Gibbs and Shannon entropy are quite similar, however eq.(2) implies that S_S is a coarse grained version of S_G where the number of internal microstates in each logic state is ignored. The relation between Gibbs and Shannon entropies is provided by [2,3]

$$S_G = S_S + S_{cond} \quad (4)$$

¹There is no unique way to make such division and the particular choice for $\Omega_0 \dots \Omega_{n-1}$ is application dependent.

^(a)E-mail: davide.chiuchiu@nipslab.org

where S_{cond} is the contribution of the different microstates inside each Ω_i .

Computation can be seen as an information manipulation performed through a physical transformation of the system. This physical transformation obeys Clausius theorem, $Q \geq -T\Delta S_G$, where Q is the amount of heat exchanged with the reservoir². Using eq.(4) we obtain

$$Q \geq -T\Delta S_S - T\Delta S_{cond}. \quad (5)$$

Eq.(5) is the generalized Landauer principle [2] where we can recognize a minimum heat production due to information processing ($-T\Delta S_S$) that can be increased or compensated by the entropy change inside each logic state ($-T\Delta S_{cond}$).

Up to now, $\Delta S_{cond} = 0$ was explicitly or implicitly assumed for any logic operation performed on bistable symmetric systems [1, 8–12], while $\Delta S_{cond} \neq 0$ was considered possible only for asymmetric systems [2]. In this paper we show that $\Delta S_{cond} \neq 0$ is also valid for binary symmetric systems and that it contributes significantly to minimum heat production. To prove this fact, we consider a bistable symmetric system and a continuous two-peaked PDF that generalizes the one used in [2, 10–13]. With this PDF we show that ΔS_{cond} can be expressed as the sum of three different terms connected to the PDF structure. The contribution of the three terms to conditional entropy (and thus to minimum heat production) is then discussed for the case of bit-reset.

II. Conditional entropy as the sum of three contributions. – We assume that our microstates space $\Omega =]-\infty, \infty[$ is split in two subsets $\Omega_0 =]-\infty, 0[$ and $\Omega_1 =]0, \infty[$. The physical system is said to encode one bit of information in the 0 or 1 logic state if $x \in \Omega_0$ or $x \in \Omega_1$, respectively. Probabilities of being in the 0 or 1 logic state and Shannon entropy are given by eq.(2) and eq.(3) with $n = 2$ and $i = 0, 1$.

To practically confine x values within a given working range we introduce a static bistable and symmetric potential $U(x)$ with two minima in x_a and x_b separated by an energy barrier ΔU in $x = 0$. Since the boundary between Ω_0 and Ω_1 lies at the top of the energy barrier, x_a and x_b are contained in different logic subsets and the energy barrier guarantees state stability for a time shorter than the residence time [14].

If the system is at thermal equilibrium with the reservoir, the associated PDF is given by the canonical distribution [15]:

$$P(x) = \mathcal{Z}^{-1} e^{-\frac{U(x)}{K_b T}} \quad (6)$$

where \mathcal{Z}^{-1} is the partition function. Eq.(6) describes a system with equal probability 1/2 of being either in the logical state 0 or in the logical states 1. The corresponding non-equilibrium distribution is usually [2, 11, 12] represented as:

$$P(x) = P_0 2\mathcal{Z}^{-1} e^{-\frac{U(x)}{K_b T}} \Theta(x) + P_1 2\mathcal{Z}^{-1} e^{-\frac{U(x)}{K_b T}} \Theta(-x) \quad (7)$$

where $\Theta(x)$ is Heaviside step function, with P_0 and P_1 probabilities of being in the 0 and 1 state, arbitrarily adjusted. Results concerning Landauer limit and conditional entropy are then derived from eq.(7).

In this paper we assume a different PDF, that includes eq.(7) as a special case, but avoids its discontinuity when P_0 and P_1 differ from 1/2. This is not just a mathematical nuisance: for any bistable system that left to reach a local equilibrium condition near its minima, the Fokker-Planck equation [16, 17] provides continuous $P(x)$ for continuous $U(x)$. Most importantly, our approach, at difference with previous results [2], shows that ΔS_{cond} can be different from 0 even for symmetric bistable potentials if the height of the energy barrier becomes comparable with the thermal noise ($\Delta U \approx K_b T$).

To identify a more appropriate form of the PDF, we note that any bistable system with a continuous potential and a moderate noise intensity, will generally have $P(x)$ peaked around x_a and x_b and suppressed near the barrier. Based on these considerations we introduce two functions, $\eta_a(x)$ and $\eta_b(x)$, single-peaked with maximum approximately at x_a and x_b respectively. Their supports are labeled Ω_a and Ω_b and those may not coincide with Ω_0 and Ω_1 . This additional freedom in Ω_a and Ω_b gives the advantage to choose η_a and η_b without discontinuities of the first kind as in eq.(7). To add to the generality, we allow Ω_a and Ω_b to superimpose over a subset called Ω_{ov} . We finally assume normalized functions $\int_{\Omega_{a,b}} \eta_{a,b}(x) dx = 1$. In this way we can express the nonequilibrium PDF as:

$$P(x) = P_a \eta_a(x) + P_b \eta_b(x). \quad (8)$$

with

$$P_a + P_b = 1. \quad (9)$$

where P_a (P_b) describe the average probability that one microstate belongs to η_a (η_b). Application of eq.(2) to eq.(8) gives

$$\begin{aligned} P_0 &= P_a \int_{\Omega_a \cap \Omega_0} \eta_a(x) dx + P_b \int_{\Omega_b \cap \Omega_0} \eta_b(x) dx \\ P_1 &= P_a \int_{\Omega_a \cap \Omega_1} \eta_a(x) dx + P_b \int_{\Omega_b \cap \Omega_1} \eta_b(x) dx \end{aligned} \quad (10)$$

showing a linear relationship between P_0 , P_1 , P_a and P_b . We stress that eq.(8), (9) and (10) are more general than eq.(7) and are valid even if $U(x)$ is simply bistable but not symmetric.

We now calculate Gibbs and conditional entropy using

²By convention, Q is positive for heat given to the reservoir

the PDF given in eq.(8). Gibbs entropy is given by:

$$\begin{aligned}
 \frac{S_G}{K_b} &= - \int_{\Omega} (P_a \eta_a + P_b \eta_b) \ln (P_a \eta_a + P_b \eta_b) dx = \\
 &= - \int_{\Omega_a \setminus \Omega_{ov}} P_a \eta_a \ln (P_a \eta_a) dx - \int_{\Omega_b \setminus \Omega_{ov}} P_b \eta_b \ln (P_b \eta_b) dx + \\
 &\quad - \int_{\Omega_{ov}} P_a \eta_a \ln \left(P_a \eta_a \left(1 + \frac{P_b \eta_b}{P_a \eta_a} \right) \right) dx + \\
 &\quad - \int_{\Omega_{ov}} P_b \eta_b \ln \left(P_b \eta_b \left(1 + \frac{P_a \eta_a}{P_b \eta_b} \right) \right) dx = \\
 &= - \int_{\Omega_a} P_a \eta_a \ln (P_a \eta_a) dx - \int_{\Omega_b} P_b \eta_b \ln (P_b \eta_b) dx + \\
 &\quad + \frac{P_a}{K_b} I(\eta_a, \eta_b, \frac{P_b}{P_a}) + \frac{P_b}{K_b} I(\eta_b, \eta_a, \frac{P_a}{P_b})
 \end{aligned} \tag{11}$$

with

$$I(\eta_a, \eta_b, q) = -K_b \int_{\Omega_{ov}} \eta_a \ln \left(1 + q \frac{\eta_b}{\eta_a} \right) dx \tag{12}$$

After some additional calculations and the definition of Gibbs single-peak entropy $S_a = -K_b \int_{\Omega_a} \eta_a \ln \eta_a dx$, S_G can be expressed as the sum of three terms:

$$S_G = S_{cg} + S_{pe} + S_{ov} \tag{13}$$

with

$$S_{cg} = -K_b (P_a \ln P_a + P_b \ln P_b) \tag{14a}$$

$$S_{pe} = P_a S_a + P_b S_b \tag{14b}$$

$$S_{ov} = P_a I(\eta_a, \eta_b, \frac{P_b}{P_a}) + P_b I(\eta_b, \eta_a, \frac{P_a}{P_b}). \tag{14c}$$

The first is S_{cg} , a coarse grained entropy. It is the Shannon entropy built with probabilities P_a, P_b that a microstate belongs to η_a and η_b . S_{pe} describes entropy contributions arising from the shapes of the two peaks when those are considered non-overlapped. Corrections due to overlap are given in S_{ov} by the I integrals.

Conditional entropy S_{cond} is then given as three contributions as

$$S_{cond} = S_G - S_S = S_{ex} + S_{pe} + S_{ov} \tag{15}$$

with

$$S_{ex} = S_{cg} - S_S. \tag{16}$$

Here S_{pe} and S_{ov} are the same as in eq.(13). The remaining term, S_{ex} , is the difference between the entropy eq.(14a) (describing microstate coarse graining in two peaks) and Shannon entropy (describing microstate coarse graining in two logic subsets). It thus represents the absolute entropic measure of the error we commit if we exchange probability set (P_a, P_b) with (P_0, P_1) . Conditional entropy variation is given by:

$$\Delta S_{cond} = \Delta S_{ex} + \Delta S_{pe} + \Delta S_{ov}. \tag{17}$$

Here ΔS_{ex} takes into account the possible change of the value of P_a and P_b , and the fact that peaks can move from one logic subset to the other during the transformation (this reflects the change in the overlap between Ω_a and Ω_b with Ω_0 and Ω_1). ΔS_{pe} arises from the changes in shape of the peaks, while ΔS_{ov} from the change in the way the two peaks overlap.

In the next sections we consider two examples to better illustrate eq.17.

III. First example. — As first example, we recall that eq.(7) is eq.(8) with

$$\eta_a = 2Z^{-1} \exp(-U(x)/(K_b T)) \Theta(x) \tag{18a}$$

$$\eta_b = 2Z^{-1} \exp(-U(x)/(K_b T)) \Theta(-x) \tag{18b}$$

$$\Omega_a = \Omega_0, \quad \Omega_b = \Omega_1 \tag{18c}$$

Conditional entropy is easily computed. Eq.(18c) and eq.(10) gives that $P_0 = P_a$ and $P_1 = P_b$ so $S_{cg} = 0$. Also $S_{ov} = 0$, because $\Omega_{ov} = \Omega_0 \cap \Omega_1 = \emptyset$. As a consequence

$$S_{cond} = S_{pe} = -K_b \ln 2 + K_b \left(\frac{\langle U \rangle}{K_b T} + \ln Z^{-1} \right) \tag{19}$$

where $\langle U \rangle$ is the average value of $U(x)$ at thermal equilibrium.

Now, when eq.(7) is used as nonequilibrium PDF [10–12], logic operations are isothermal transformations that change (P_0, P_1) from some initial values (P_0^i, P_1^i) into some others (P_0^f, P_1^f) . As a consequence $\Delta S_{cond} = 0$ for any logic operation. However, this is not a general property of symmetric bistable systems [2] but a consequence of the assumption of eq.(7) as nonequilibrium PDF. We illustrate this point with the second example.

IV. Gaussian example. — Let us consider the system with symmetric bistable potential $U(x)$ at thermal equilibrium with a reservoir at a temperature T . Due to the symmetry $\eta_b(x) = \eta_a(-x)$ and

$$P_a^i = P_b^i = \frac{1}{2}. \tag{20}$$

From eq.(10) we have

$$P_0^i = P_1^i = \frac{1}{2} \tag{21}$$

i.e. the initial state of the bit is undefined.

A simple analytical expression for $\eta_a(x)$ can be obtained through harmonic approximation of $U(x)$ near x_a minimum in eq.(6):

$$\eta_a(x) \approx \mathcal{N} e^{-\frac{(x+x_a)^2}{2\sigma^2}} \tag{22}$$

with σ fitted from local harmonic approximation

$$\sigma = \sqrt{\frac{K_b T}{U''(x_a)}}, \tag{23}$$

U'' the second derivative respect to x , $\mathcal{N} = (\sqrt{2\pi}\sigma)^{-1}$ and

$$h = \frac{|x_a|}{\sigma} = \sqrt{2} \sqrt{\frac{\frac{1}{2}U''(x_a)x_a^2}{K_bT}}. \quad (24)$$

The numerator of eq.(24) is the harmonically approximated potential energy of a point distant $|x_a|$ from x_a . If $U(x)$ is smooth, this quantity has approximately the same amplitude of ΔU , implying $h \approx [2\Delta U/(K_bT)]^{\frac{1}{2}}$ and

$$h \gg 1 \equiv \frac{\Delta U}{K_bT} \gg 1 \quad (25)$$

A finite-time protocol is then implemented to reset the logic state to zero, i.e. the average value of x becomes negative. The protocol may be the same given in [13], but any physical transformation that recovers $U(x)$ at the end and that makes the average value of x negative is indeed a valid choice. If both $U(x)$ and the protocol are specified, then the final PDF is solution of the associated Fokker-Planck equation [16, 17]. The main limitation of this approach is that analytical solutions are difficult to obtain. For this reason we assume that at the end of the protocol the system is allowed to reach a local equilibrium condition around the minima of $U(x)$. The final PDF is then correctly approximated by eq.(8) with $\eta_b(x) = \eta_a(-x)$ and η_a , σ and h given by eq.(22), (23) and (24). The sole difference with the initial distribution is that P_a^f can take any value in $[0.5, 1]$ so to have negative x average value. In this way we can also study reset protocols where there is a non-negligible probability P_1^f to end in the wrong logic state. Putting eq.(22), $\eta_b(x) = \eta_a(-x)$ and $\Omega_a = \Omega_b = \Omega =]-\infty, \infty[$ in eq.(10), after some algebraic manipulations we obtain

$$\begin{aligned} P_0^f &= \frac{1 + (2P_a^f - 1) \operatorname{erf}(\frac{h}{\sqrt{\pi}})}{2} \\ P_1^f &= \frac{1 + (2P_b^f - 1) \operatorname{erf}(\frac{h}{\sqrt{\pi}})}{2}. \end{aligned} \quad (26)$$

where $\operatorname{erf}(y) = 2\pi^{-\frac{1}{2}} \int_0^y e^{-z^2} dz$ is the error function.

Summarizing, we have that our bit-encoding system is initially at thermal equilibrium: its initial PDF (Fig. 1, left) is $P(x) = \eta_a(x)/2 + \eta_a(-x)/2$ with $\eta_a(x)$ given by eq.(22). Logic states probabilities are given by eq.(21). We then perform a finite-time reset to zero protocol which brings the system in the final PDF (Fig. 1, right) described by eq.(8) with $\eta_b(x) = \eta_a(-x)$, $\eta_a(x)$ given by eq.(22) and $P_a^f \in [0.5, 1]$. Logic state state probabilities are given by eq.(26). The hypothesis here used are common in experiments on bit reset [8, 13, 18] with the sole new addition that initial and final probability density functions have the form of eq.(8). This choice is justified for reset protocols which brings the system at least in local equilibrium near potential minima. The possibility of reset errors is also included in this description.

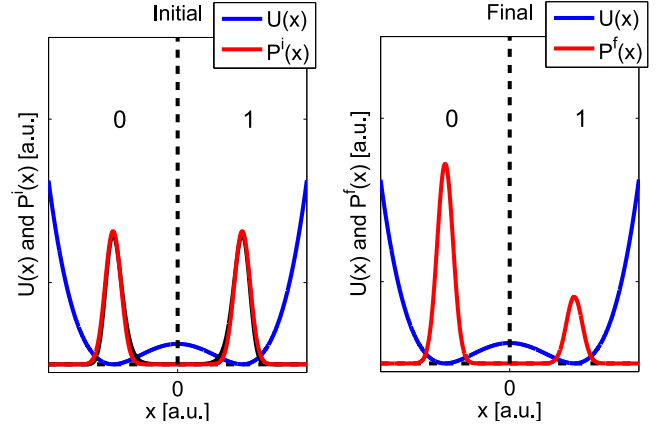


Figure 1: Schematic representation of a sample reset protocol. Left and right plots describe the initial and final state of the bit. Potential $U(x)$ (blue) is the same at the begin and the end of the protocol. Initial and final distributions are shown in red. Final distribution has a negative average value of x , so the bit is on the average reset to zero.

Next step is the computation of the entropy variations from their definitions with the Gaussian peaks. Gibbs entropy, ΔS_G is given by:

$$\begin{aligned} \frac{\Delta S_G}{K_b} &= -P_a^f \ln P_a^f - P_b^f \ln P_b^f - \ln 2 \\ &\quad - \frac{P_a^f}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} \ln \left(1 + \frac{P_b^f}{P_a^f} e^{2zh-2h^2} \right) dz + \\ &\quad - \frac{P_b^f}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} \ln \left(1 + \frac{P_a^f}{P_b^f} e^{2zh-2h^2} \right) dz + \\ &\quad + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} \ln \left(1 + e^{2zh-2h^2} \right) dz \end{aligned} \quad (27)$$

with $z = (x \pm h\sigma)/\sigma$. The first line of the RHS represents ΔS_{cg} , while the last three lines are ΔS_{ov} . $\Delta S_{pe} = 0$ here because eq.(9) is satisfied and $\eta_a(x)$ is the same at the beginning and the end of the protocol.

Shannon entropy variation is promptly obtained through application of eq.(21) and (26) to eq.(3)

$$\begin{aligned} \frac{\Delta S_S}{K_b} &= - \frac{1 + (2P_a^f - 1) \operatorname{erf}(\frac{h}{\sqrt{\pi}})}{2} \ln \frac{1 + (2P_a^f - 1) \operatorname{erf}(\frac{h}{\sqrt{\pi}})}{2} + \\ &\quad - \frac{1 + (2P_b^f - 1) \operatorname{erf}(\frac{h}{\sqrt{\pi}})}{2} \ln \frac{1 + (2P_b^f - 1) \operatorname{erf}(\frac{h}{\sqrt{\pi}})}{2} + \\ &\quad - \ln 2. \end{aligned} \quad (28)$$

Using now eq.(4) and (17) we obtain

$$\begin{aligned} \frac{\Delta S_{cond}}{K_b} = & -P_a^f \ln P_a^f - P_b^f \ln P_b^f + \\ & + \frac{1 + (2P_a^f - 1) \operatorname{erf}(\frac{h}{\sqrt{\pi}})}{2} \ln \frac{1 + (2P_a^f - 1) \operatorname{erf}(\frac{h}{\sqrt{\pi}})}{2} + \\ & + \frac{1 + (2P_b^f - 1) \operatorname{erf}(\frac{h}{\sqrt{\pi}})}{2} \ln \frac{1 + (2P_b^f - 1) \operatorname{erf}(\frac{h}{\sqrt{\pi}})}{2} + \\ & - \frac{P_a^f}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} \ln \left(1 + \frac{P_b^f}{P_a^f} e^{2zh - 2h^2} \right) dz + \\ & - \frac{P_b^f}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} \ln \left(1 + \frac{P_a^f}{P_b^f} e^{2zh - 2h^2} \right) dz + \\ & + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} \ln \left(1 + e^{2zh - 2h^2} \right) dz. \end{aligned} \quad (29)$$

where the first three lines of the RHS are ΔS_{ex} .

Eq.(27) and (29) are plotted in Fig. 2 as functions of P_a^f and h . Figure 2.1 shows our main result, that $\Delta S_{cond} \neq 0$ even for a system with a symmetric bistable potential $U(x)$. For this particular example, $\Delta S_{cond} \leq 0$ with $-0.12K_b \ln 2$ as largest deviation from zero. Non-zero values are obtained for $h \lesssim 1$ and $P_a^f \approx 1$, a behaviour that can be explained in terms of eq.(17), eq.(23) and eq.(25). P_a^f must be close to one because, for $P_a^f \approx 0.5$, no reset takes place and entropy variations are zero. Moreover, for $h \lesssim 1$, thermal noise intensity $K_b T$ is comparable or greater than ΔU and $\sigma \gtrsim |x_a|$; this implies that each Gaussian peak has from 50% to 85% of its total area inside one logic subset and the remaining part in the other. In this case we see from eq.10 that $(P_0^f, P_1^f) \neq (P_a^f, P_b^f)$ by a significative amount. Thus, thermal noise limits the identification of a single logic state with a well defined Gaussian peak, providing $\Delta S_{ex} \neq 0$. The discussion for $\Delta S_{ov} \neq 0$ is similar but less intuitive. When the noise is large ($h \approx 1$), the two Gaussian peaks overlap within 1σ . Peaks overlap is thus a relevant property of $P(x)$ and the integrals $I(\cdot)$ in eq.(14c) must be nonzero. Additionally, their sum in S_{ov} strongly depends on (P_a, P_b) specific values, so also ΔS_{ov} must be nonzero.

Since ΔS_{cond} can be nonzero for a reset protocol operated on a bistable symmetric system, we investigate its contributions to minimum heat production. This is given by

$$Q_{min} = -T\Delta S_G = -T\Delta S_S - T\Delta S_{cond} \quad (30)$$

where $-T\Delta S_S$ is the heat production for a reset with error probability P_1^f [13, 19] and $-T\Delta S_{cond}$ is the contribution due to conditional entropy. In our example we proved that $0 \leq -T\Delta S_G \leq K_b T \ln 2$ (Fig.2.2) and that $-T\Delta S_S \leq -T\Delta S_G$ (Fig.2.1). Putting everything together, we obtain

$$0 \leq -T\Delta S_S \leq Q_{min} \leq K_b T \ln 2. \quad (31)$$

Two results are drawn from these inequalities. The first one is the well known $0 \leq -T\Delta S_S \leq K_b T \ln 2$ and states that error probabilities reduce the minimum heat produced to reset one bit [13, 19]. The second is $-T\Delta S_S \leq$

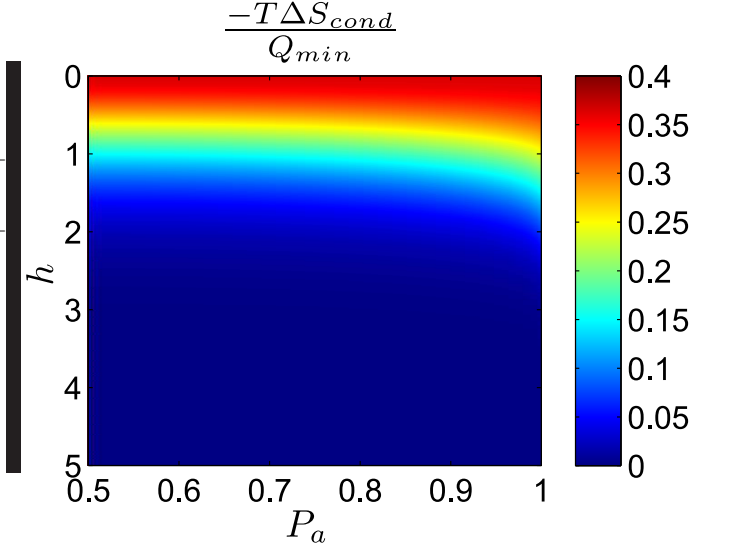


Figure 3: The relative contribution of $-T\Delta S_{cond}$ to Q_{min} as a function of P_a^f and h . From this plot it's clear that ΔS_{cond} is non-negligible if $h \lesssim 1$.

Q_{min} and tells us that, in our example, minimum heat production is underestimated if ΔS_{cond} is improperly neglected. For a quantitative evaluation of the underestimation, see Fig. 3

To conclude, we discuss the low noise limit $h \gg 1$ (eq.(25)). From eq.(26), (27), (28) and (29) we have:

$$P_0^f \approx P_a^f, \quad P_1^f \approx P_b^f \quad (32a)$$

$$\Delta S_G \approx \Delta S_{cg} = -K_b(P_a^f \ln P_a^f + P_b^f \ln P_b^f + \ln 2) \quad (32b)$$

$$\Delta S_S \approx \Delta S_{cg} = -K_b(P_a^f \ln P_a^f + P_b^f \ln P_b^f + \ln 2) \quad (32c)$$

$$\Delta S_{cond} \approx 0. \quad (32d)$$

Putting eq.(32a), (32c) and (32d) in eq.(5), we obtain

$$Q \geq K_b T (P_0^f \ln P_0^f + P_1^f \ln P_1^f + \ln 2) \quad (33)$$

which is the classical Landauer limit for bit-reset with error probability P_1^f [10–13, 19]. This occurs because both the differences between (P_a, P_b) and (P_0, P_1) and the overlap integrals are given by Gaussian tails. For this reason they can be neglected and $\Delta S_{ex} \approx 0$, $\Delta S_{ov} \approx 0$ and $\Delta S_{cond} \approx 0$.

In the literature [10–13, 19], eq.(33) is obtained with eq.(7) as nonequilibrium PDF. In this paper we derive this result using weaker (and physically more reasonable) assumptions: eq.(8) as nonequilibrium PDF and $\Delta U/(K_b T) \gg 1$. If we remove $\Delta U/(K_b T) \gg 1$ assumption we obtain again eq.(31), where the contribution of the conditional entropy to Q_{min} becomes significant. As proved in the previous example, such contributions can't arise from eq.(7) if $\Delta U/(K_b T) \gg 1$ is removed.

V. Conclusions. – To study the role of conditional entropy in bistable symmetric systems, we proposed eq.(8) as a new way to represent nonequilibrium PDFs instead

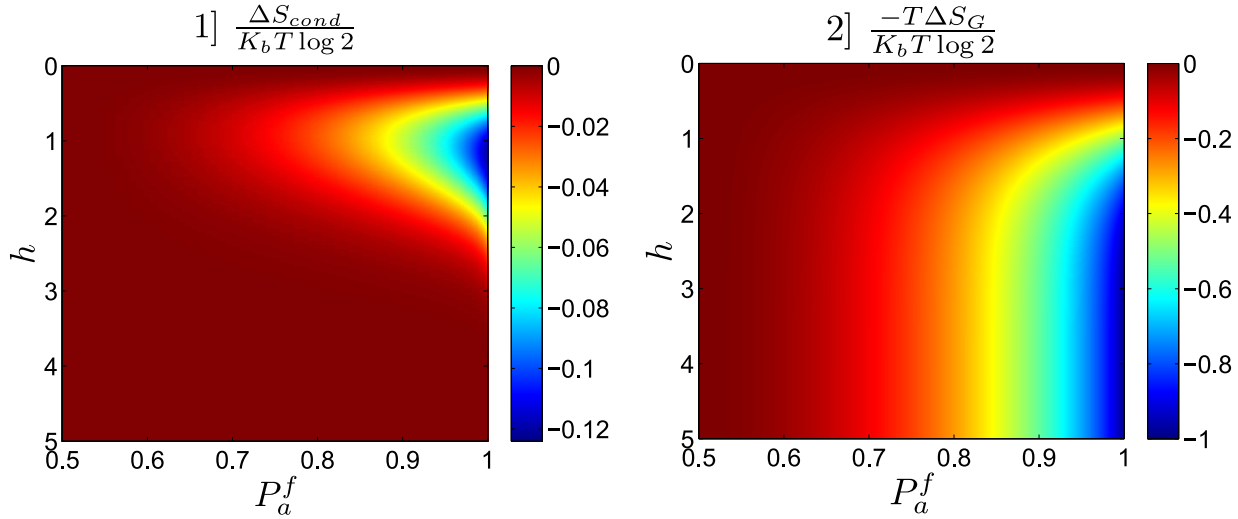


Figure 2: Entropy variations for faulty reset protocols as functions of P_a^f and h . The latter spans in $[0, 5]$ because $h = 5$ suffices to reach the asymptotic limit of small noise values (see eq.(27) and (29) when $h \gg 1$). In order: 1) $\Delta S_{cond}/(K_b \ln 2)$, 2) $\Delta S_G/(K_b \ln 2)$.

of the commonly used eq.(7). This more general PDF, which includes eq.(7) as special case, allowed us to show that conditional entropy plays a role also for symmetric bistable systems whereas the formulation based on eq.(7) always gives a null contribution. Thanks to this new formulation the conditional entropy can be written as the sum of three contributions directly and intuitively connected to the PDF structure.

To illustrate the aforementioned points, we discussed the reset of one bit with eq.8 as nonequilibrium PDF. Here two scenarios are available. If $\Delta U/(K_b T) \gg 1$, then $\Delta S_{cond} = 0$ and known results on Landauer principle are obtained with a weaker set of assumptions. If $\Delta U/(K_b T) \lesssim 1$, we obtain a new result: ΔS_{cond} is nonzero for bistable and symmetric systems and contributes significantly to the minimum heat production. We underline that the latter scenario is not of academical interest only: the scaling down trend in ICT is likely to produce a device that operate in the $\Delta U/(K_b T) \lesssim 1$ regime within few years.

We acknowledge support by the European Union (FPVII (2007-20013) under G.A. n 318287 LANDAUER, G.A. n 270005 ZEROPOWER and G.A. n 611004 ICT-Energy).

References

- [1] R. Landauer. Irreversibility and heat generation in the computing process. *IBM J. Res. Develop.*, 5:183–191, 1961.
- [2] T. Sagawa. Thermodynamic and logical reversibilities revisited. *J. Stat. Mech.*, 2014.
- [3] O. J. E. Maroney. Generalizing landauer’s principle. *Phys rev E*, 79(3), 2009.
- [4] T.M. Cover and J.A. Thomas. *Elements of Information Theory*. John Wiley and Sons, 1991.
- [5] E. T. Jaynes. Gibbs vs boltzmann entropies. *AJP*, 33(5):391–398, 1965.
- [6] U. Seifert. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.*, 75:126001, 2012.
- [7] C. E. Shannon. *The Mathematical Theory of Communication*. University of Illinois Press, 1949.
- [8] B. Lambson and J. Bokor D. Carlton. Exploring the thermodynamic limits of computation in integrated systems: Magnetic memory, nanomagnetic logic and the landauer limit. *Phys. Rev. Lett.*, 107:010604, 2011.
- [9] S. Deffner and C. Jarzynski. Information processing and the second law of thermodynamics: an inclusive hamiltonian approach. *Phys. Rev. X*, 3:041003, 2013.
- [10] K. Shizume. Heat generation required by information erasure. *Phys. Rev. E*, 52(4):3495–3499, 1995.
- [11] B. Piechocinska. Information erasure. *Phys. Rev. A*, 61:062314, 2000.
- [12] R. Dillenschneider and E. Lutz. Memory erasure in small systems. *PRL*, 102:210601, 2009.
- [13] A. Berut, A. Arakelyan, A. Petrosyan, S. Ciliberto, R. Dillenschneider, and E. Lutz. Experimental verification of landauers principle linking information and thermodynamics. *Nature*, 483:187–189, 2012.
- [14] L. Gammaitoni, P. Hnggi, P. Jung, and F. Marcheson. Stochastic resonance. *Rev. Mod. Phys.*, 70:223–287, 1998.
- [15] C. M. Van Vliet. *Equilibrium and Non-equilibrium Statistical Mechanics*. World Scientific Publishing, 2008.
- [16] C. Gardiner. *Stochastic Methods. A Handbook for the Natural and Social Sciences*. Springer, fourth edition, 2009.
- [17] H. Risken. *The Fokker-Planck equation. Methods of solution and Applications*. Springer, second edition, 1989.
- [18] A. Berut, A. Petrosyan, and S. Ciliberto. Detailed jarzynski equality applied on a logically irreversible procedure. *EPL*, 103(6), 2013.
- [19] L. Gammaitoni. Beating the landauer’s limit by trading

energy with uncertainty. *pre-print*, 2011. arXiv:1111.2937.